Journal of Experimental Psychology: Applied
2004, Vol. 10, No. 4, 238–244

1076-898X/04/$12.00 DOI: 10.1037/1076-898X.10.4.238

Frames of Reference in Mobile Augmented Reality Displays

Weimin Mou
Chinese Academy of Sciences

Frank Biocca, Charles B. Owen, Arthur Tang,
Fan Xiao, and Lynette Lim
Michigan State University

In 3 experiments, the authors investigated spatial updating in augmented reality environments. Participants learned locations of virtual objects on the physical floor. They were turned to appropriate facing directions while blindfolded before making pointing judgments (e.g., “Imagine you are facing X. Point to Y”). Experiments manipulated the angular difference between the learning heading and the imagined heading and between the actual heading and the imagined heading. The effect of actual–imagined on pointing latency was observed for naïve users but not for users with brief training or instructions concerning the fact that objects can move with body movements. The results indicated that naïve users used an environment-stabilized reference frame to access information arrays, but with experience and instruction the nature of the representation changed from an environment stabilized to a body stabilized reference frame.

The term *augmented reality* (AR) describes systems that blend computer-generated virtual objects or environments with real environments (Azuma, 1997; Barfield & Caudell, 2000). In a typical AR system for augmented vision, a see-through head-mounted display (HMD) is used to overlay computer-generated graphics on the real environment in real time. The advent of stereo AR head-mounted displays and tracking technologies allows user-interface designers to array information spatially around the body of a mobile computer user, as is illustrated in Figure 1. Information can be presented in the form of two-dimensional (2D) pages and/or 3D objects at any spatial location relative to the user. Typically, virtual objects are registered to the real world, giving the appearance that objects are floating in space and attached to some invisible reference coordinate frame. Many such frames are possible including frames stabilized relative to the environment, objects in the environment, the head, the torso, or any extremities. Generally, those frames could be divided into two categories: one stabilized with the user’s body and one stabilized with the physical environment.

It is clear that each frame of reference has its own advantage in some typical applications and placement of some type of information over the other. However, it is not clear which frame of reference matches the nature of human spatial memory and spatial updating of the information array better, or how to manipulate a user’s preference of frames of reference according to different applications. Clearly, both frames of reference have physical analogs with which users are familiar. Data in an environment-stabilized frame of reference are analogous to objects physically located in the environment. Data in a body-stabilized frame of reference are analogous to tools worn on the belts and clothing. However, physical experience does not prepare users for data presentation where 2D and 3D data objects appear to hover weightlessly in the frame of reference, or where the data can freely move. AR makes these new concepts possible and relatively easy to implement.

The first question that guided the research reported here can be put in a simple way: What is the default frame of reference for information arrays in the AR environments? For example, when users turn left with their eyes closed, will users expect the surrounding information arrays to move with their body or should they stay still with respect to the world?

There has been no previous research into human spatial memory and spatial updating of information array in mobile AR systems. However, some answers to the above questions may be suggested by human spatial memory and spatial updating of real objects in the physical world. There is a large body of evidence indicating that people update locations of objects during locomotion (e.g., Farrell & Robertson, 1998; Mou, McNamara, Valiquette, & Rump, 2004; Presson & Montello, 1994; Rieser, 1989; Rieser, Guth, & Hill, 1986; Sholl & Bartels, 2002; Simons & Wang, 1998; Waller, Montello, Richardson, & Hegarty, 2002; Wang & Simons, 1999). For example, participants in one of Waller et al.’s (2002) experiments learned 4-point paths. In the “stay” condition, participants remained at the study position and made pointing judgments from headings of 0° and 180° (“aligned” vs. “misaligned”). The results in this condition replicated several other studies of spatial memory in showing that performance was better for the imagined heading of 0° than for the imagined heading of 180° (e.g., Levine, Jan-

---

Weimin Mou, Institute of Psychology, Chinese Academy of Sciences, Beijing, China. Frank Biocca and Lynette Lim, Department of Telecommunications, Michigan State University.Charles B. Owen, Arthur Tang, and Fan Xiao, Department of Computer Science, Michigan State University.

Preparation of this article and the research reported herein were supported in part by the National Science Foundation (NSF) Grant IIS 0222831, and the “One-Hundred-Talent Plan” of the Chinese Academy of Sciences. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

Correspondence concerning this article should be addressed to Weimin Mou, Institute of Psychology, Chinese Academy of Sciences, 10A Datun Road, Beijing, China 100101. E-mail: mouw@psych.ac.cn

Figure 1. An example of the mobile augmented reality system.

kovic, & Palij, 1982). In the "rotate–update" condition, participants learned the layout and then were told to turn 180° in place so that the path was behind them. Performance was now better for the heading of 180° (the new egocentric heading) than for the heading of 0° (the original learning heading). This result indicated that, as they turned, participants updated their orientation with respect to the locations in memory.

Simons and Wang (1998; also see Wang & Simons, [1999]) investigated the interaction between observer movement and layout rotations on change detection. They showed that detection of changes to a recently viewed layout of objects was disrupted when the layout was rotated to a new view and the observer remained stationary, but there was no disruption when the layout remained stationary and the observer moved to the new viewpoint. In other words, updating was efficient when the observer moved around the layout but not when the layout rotated in front of the observer (see Wraga, Creem, & Profitt, 2000, for analogous results in imagined updating).

In a recent study, Mou, McNamara, et al. (2004) reported that the angular distance between both the imagined heading and the learning heading and the imagined heading and the actual heading had effects on people's ability to accurately point to objects in the environment. Participants in one of their experiments learned the locations of 10 objects from a single view (e.g., a vase was located next to the learning position; see Figure 7 of that article), walked to the center of the layout (e.g., next to a shoe), and faced three headings before making pointing judgments from imagined headings. There were three imagined headings, 0° (e.g., "Imagine you are facing the phone), 90° ("Imagine you are facing the banana"), or 225° ("Imagine you are facing the jar"), and there were two angular distances between the imagined heading and the actual heading, 0° (e.g., participants actually faced the phone and were instructed to imagine facing the phone) or 225° (e.g., participants actually faced the book and were instructed to imagine facing the phone). Pointing performance was best when the imagined heading was parallel to the learning view. Pointing performance was also better when the actual and the imagined headings were the same. Mou, McNamara, et al. proposed that people both represent locations of objects in terms of an object-to-object frame of reference selected by the egocentric view (also see Mou & McNamara, 2002; Mou, Zhang, & McNamara, 2004) and update their location and orientation in terms of that frame of reference during locomotion.

In Experiment 1, using the paradigm developed by Mou, McNamara, et al. (2004), we investigated whether people with no experience in mobile AR systems would use the environment-stabilized or body-stabilized frame of reference as the default. We hypothesized that if participants used the body-stabilized frame of reference, the angular distance between the imagined heading and the actual heading would not affect pointing performance; however, if they used the environment-stabilized frame of reference, the angular distance between the imagined heading and the actual heading would affect pointing performance, just as was observed in the Mou, McNamara, et al. study.

We only investigated the user's frame of reference preference during rotation (and only in the horizontal plane) rather than in translation (in all three body axes); we assumed the information objects around the user's body should be arrayed independently of the user's translation. We limited our study to the frame of reference preference during body rotation rather than head rotation because a display stabilized with respect to the head would have a very limited information field.

In this study, the second goal was to examine whether the nature of the representation of the objects in the AR system can be altered from environment centered to body centered. The experience of large objects moving with the body does not occur normally in the real world, except in cases in which objects are directly attached to the body. So, although the default organization of virtual objects appears to be tied to the exocentric world frame, experience of a body-stabilized frame might enable users to adopt the newly experienced frame when updating their memories for objects' locations in a new layout, even without direct visual guidance. In Experiment 2, we examined whether a couple of minutes of experience in the body-stabilized AR display would allow users to adopt the body-stabilized frame of reference. In Experiment 3, we examine whether only oral instructions to use a body-stabilized frame of reference for updating the location of a set of objects might be sufficient to induce participants to use a body-stabilized frame of reference in accessing a complex layout. Waller et al. (2002) reported that people were able to imagine simple, body-stabilized 4-point paths in front of them when they physically turned back after being instructed to do so.

Experiment 1

In Experiment 1, participants learned the locations of virtual objects displayed on the floor from a single stationary viewing position in a large cylindrical room; they were instructed either to keep their heading or turn 90° before making judgments of egocentric pointing (e.g., “Imagine you are facing X. Point to Z”).

Method

Participants. Participants were 8 female and 8 male undergraduates at Michigan State University who participated as partial fulfillment of course requirements.

Materials and design. Stimulus materials were displayed in stereo with the Sony Glasstron LDI-100B. Participants’ head motion was tracked with a Polhemus Fastrak magnetic tracker. Stereo graphics were rendered in real time on the basis of the data from the tracker. Presentation of stimulus materials, audio instructions for participants, experimental procedure sequencing, and data collection for the experiment were automated so that the experimenter did not need to hand code the experimental results. The program was written with ImageTclAR (Owen, Tang, & Xiao, 2003).

A configuration of eight virtual objects was displayed by the AR system on the floor approximately 1.4 m from the observer (see Figure 2). Objects were selected with the restrictions that they be visually distinct, fit within an area approximately 0.3 m on each side, and not share any obvious semantic associations. The objects were all virtual analogs of existing physical objects and were presented in exact scale.

Each test trial was constructed from the names of two objects in the layout and required participants to point to an object (e.g., “Imagine you are facing the cell phone; please point to the ball”). The first object established the imagined heading (e.g., cell phone) and the second object was the target (e.g., ball). Participants pointed with a hand-held wand.

The design is illustrated in Figure 3. The independent variables were (a) the angular difference between the learning heading and the imagined heading at the time of test and (b) the angular difference between the actual body heading and the imagined heading at the time of test. As is shown in Figure 2, to factorially manipulate these two variables, participants had two actual body headings at the time of test: One was the same as the learning heading (e.g., actually facing the cell phone), and the other was 90° different from the learning heading (e.g., actually facing the book). At each actual heading, participants had two imagined headings: One was the same as the learning view (e.g., “Imagine you are facing the cell phone”), and the other was 90° from the learning view (e.g., “Imagine you are facing the book”). Hence, as illustrated in Figure 3, the actual body heading was the same as the learning heading when the distances of the learning–imagined and the actual–imagined were the same (both were 0° or 90°), or it was 90° different from the learning heading when the distances of the learning–imagined and the actual–imagined were different (one was 0° and the other was 90°). Both of these variables were manipulated within participants. At each actual body heading, participants had 14 trials (pointing to each of the seven objects, except the imagined facing object at each imagined heading) in a random order. Participants would imagine themselves or the scene rotating 90° when the imagined heading was 90° from their actual heading (e.g., they believed they were actually facing the cell phone but were required to imagine facing the book). According to the Wraga et al. (2000) study, most people would rotate their body. However, in this study, we did not explicitly instruct participants to adopt body or scene rotation when the imagined heading was different from the actual heading because that is beyond our focus and would not change our findings.

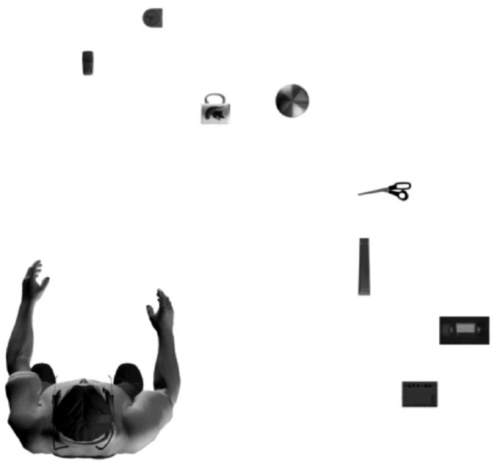

Figure 2. Layout of objects used in the present experiments. (During the learning phase, half of the participants faced the cell phone and the other half faced the notebook)

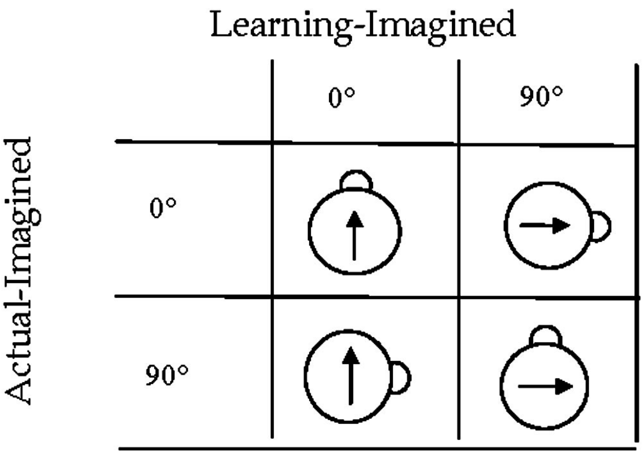

Figure 3. Design of the experiments. Head–nose icons indicate actual headings; arrows indicate imagined headings. Learning headings are not indicated in the figure but were always consistent with the right side up direction of the page. To maintain consistency with previous experiments, headings and differences between them were measured counterclockwise with respect to the right side up direction of the page.

During the learning phase, half of the participants were randomly assigned to face the cell phone and the other half faced the book. This design counterbalanced the pointing direction across all four conditions (as is illustrated in Figure 3) and ensured that all conditions were equally difficult in terms of the pointing response. The order of participants’ actual body heading at test time was also counterbalanced across participants: Half of them kept their learning orientation in the first block of pointing and then turned 90° for the second block; the other half performed in the reverse order. The primary dependent variables were pointing latency and pointing accuracy. We calculated pointing directions in terms of the participants’ facing direction.

Procedure. Participants were randomly assigned to each body-heading combination at test time, with the constraint that each group contained an equal number of men and women.

After providing informed consent, participants were trained in how to point from the imagined heading, which is either the same as or different from their actual heading. After participants understood how to conduct the pointing judgment, the experimenter escorted them to the learning room. To remove any potential influence of environmental structures, which may represent in spatial memory, participants were blindfolded while being escorted into the learning room and to the learning position.

When the participants were standing in the learning position and facing the learning direction, the blindfold was removed. Then the participants were instrumented with the AR hardware system. The experimenter put a binder with a tracker on the participants’ waist, placed the HMD with a tracker on their head, and handed them a pointing wand with a tracker.

They were instructed to press a button on the wand when they felt they were pointing to the target object accurately. At this point, the learning phase began. Participants were instructed via earphones mounted on the HMD to point to all objects twice in a row with vision guidance (e.g., "Please point to the ball") to get used to the wand. Participants used earphones throughout the experiment to avoid any spatial references resulting from sound placement. After that, they were allowed to study the layout for 30 s, and then they were asked to keep their eyes closed and to point to the objects named by the system. Participants performed five study–test sequences and were then able to point to all of the objects accurately (within 15°). All audio cues were prerecorded in the system for consistency.

After participants had learned the layout, they were blindfolded and adopted the first actual body heading. Participants always stood at their learning position but turned their body if the actual body heading was different from the learning view. Test trials were presented and participants were asked to point with the wand as accurately as possible before they pressed the button. The tracker on the wand recorded the pointing direction; pointing latency was recorded from the onset of the target object cue to the buttonpress. After they finished all 14 trials, they adopted the second actual body heading (turned by the experimenter) and repeated the same 14 trials.

Results and Discussion

Pointing accuracy and pointing latency as a function of actual–imagined distance and learning–imagined distance are presented in Table 1. We analyzed means for each participant and for each condition through the use of repeated-measures analyses of variance (ANOVAs) with terms of actual–imagined distance (0° and 90°) and learning–imagined heading (0° and 90°).

The ANOVA results for pointing accuracy and pointing latency are presented in Table 2. In angular error, no main effect was significant. People were highly accurate in all conditions. In pointing latency, both main effects of learning–imagined and actual–imagined were significant, whereas the interaction between them was not.

The most important result of Experiment 1 was that users' pointing latency was shorter when the actual and the imagined headings were the same (0°) than when they were different (90°). This result indicates that people update the location of the virtual object when they rotate their body. In other words, humans use an environment-stabilized frame of reference to access information arrays. The evidence for this is the cost in latency that was incurred by the need to align the egocentric front with the facing object specified in the pointing judgment. Given that the imagined heading is the same, when participants use a body-stabilized frame of reference, the pointing latency should be the same when they are facing the learning view as when they are turned 90° from it. This result showed that turning 90° from the learning view at test time benefited the imagined heading of 90° but had the reverse affect on the imagined heading of 0°.

Table 1
Pointing Latency (in Seconds) and Pointing Accuracy (in Degrees) as a Function of Actual–Imagined (A–I) Distance and Learning–Imagined (L–I) Distance in Experiment 1

	A–I = 0°				A–I = 90°			
	Latency		Accuracy		Latency		Accuracy	
L–I	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
0°	3.821	2.139	14	8	4.515	2.081	15	9
90°	4.384	1.683	15	11	5.117	2.097	17	11

Table 2
Analysis of Variance Results for Pointing Latency and Pointing Accuracy in Actual–Imagined (A–I) and Learning–Imagined (L–I) Conditions in Experiment 1

	F(1, 15)		Cohen's *f*	
Source	Latency	Accuracy	Latency	Accuracy
L–I	4.52*	.32	.55	.15
Error	1.20	123.60		
A–I	11.11**	.77	.86	.23
Error	.73	39.11		
L–I × A–I	.01	.19	.00	.11
Error	1.33	39.75		

* $p < .05$. ** $p < .01$.

The second important finding was that pointing latency was shorter when the imagined and learning headings were the same (0°) than when they were different (90°). This result indicates that people represent the location of the virtual object with a frame of reference selected by the learning view; that is, spatial memory is orientation dependent.

Both of these results were also reported in the research of spatial updating of physical objects (Mou, McNamara, et al., 2004), suggesting that people code and process locations of virtual objects using the same code and process as they do for physical objects.

Experiment 2

The results of Experiment 1 indicate that participants used an environment-stabilized frame of reference to access the location of virtual objects if they had never experienced the possibility that objects can also be attached to the body (egocentric frame of reference) in virtual and AR environments.

In Experiment 2, we examined whether directly experiencing a body-stabilized display in which objects translate and rotate around the moving body (a condition rarely experienced in the physical world) would stop participants from updating their actual heading with respect to the layout but would, instead, cause them to use the body-stabilized frames of reference for other layouts. Evidence of this effect would suggest that users could learn to use and update arrays of menus and objects organized around their moving body.

Method

Participants. Participants were 8 female and 8 male undergraduates at Michigan State University who participated as partial fulfillment of course requirements.

Materials, design, and procedure. The materials, design, and procedure of Experiment 2 were similar to those of Experiment 1 except a training session was added before participants learned the experimental layout of eight objects.

At training time, five virtual objects (illustrated in Figure 4) were presented on the physical floor. Participants were instructed to look at the

Figure 4. Layout of objects used in the training session in Experiment 2.

locations of all objects. After they saw all of them, they were asked to turn left and look at the locations from the new viewing direction (note that all objects maintained their position relative to the participants' body). Then they turned back to adopt the original orientation and took a look at the locations of the objects. And then they turned right and looked at the locations from the new viewing direction (note that all objects maintained their position relative to the participants' body; that is, objects rotated when the body rotated). At last, they turned back to the original orientation. The whole training session lasted about 2 min and then the learning session started. The experimenter did not comment on or verbally explain the behavior of the virtual objects. Thus, learning about the objects was through observation only.

Results and Discussion

Pointing accuracy and pointing latency as a function of actual–imagined distance and learning–imagined distance are presented in Table 3. Means for each participant and each condition were analyzed in repeated-measures ANOVAs in terms of actual–imagined distance (0° and 90°) and learning–imagined heading (0° and 90°). The ANOVA results for pointing accuracy and pointing latency are presented in Table 4. In angular error, no effect was significant. People were highly accurate in all conditions. In pointing latency, only the main effect of learning–imagined was significant.

The most important finding of Experiment 2 was that on pointing latency the effect of the angular distance between the imagined heading and the actual heading was not significant. Although failing to reject the null hypothesis is not the same as demonstrating the validity of the null hypothesis, it is safe to conclude that the effect of the actual–imagined heading on pointing latency decreased after people had a brief exposure to a body-stabilized display. The difference in pointing latency between actual–imagined (0° and 90°) decreased from 713 ms in Experiment 1 to 394 ms in Experiment 2. The effect size *f* on pointing latency consistently decreased from .86 in Experiment 1 to .40 in Experiment 2. It is hard to exclude the possibility that some participants showed the actual–imagined effect and others did not because this is not an individual-based experiment. In general, however, the results indicate that participants were able to use the body-stabilized, egocentric frame of reference to access information for the location of an array of virtual objects after only 2 min of exposure to the location of an array of virtual objects.

Table 3
Pointing Latency (in Seconds) and Pointing Accuracy (in Degrees) as a Function of Actual–Imagined (A–I) Distance and Learning–Imagined (L–I) Distance in Experiment 2

	A–I = 0°				A–I = 90°			
	Latency		Accuracy		Latency		Accuracy	
L–I	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
0°	4.077	2.030	13	9	4.510	2.947	19	17
90°	5.513	3.347	17	7	5.777	3.441	19	12

Table 4
Analysis of Variance Results for Pointing Latency and Pointing Accuracy in Actual–Imagined (A–I) and Learning–Imagined (L–I) Conditions in Experiment 2

	F(1, 15)		Cohen's *f*	
Source	Latency	Accuracy	Latency	Accuracy
L–I	20.21**	0.26	1.16	.13
Error	1.45	77.96		
A–I	2.40	3.48	.40	.48
Error (A–I)	.81	73.91		
L–I × A–I	.05	1.07	.05	.27
Error	2.50	56.05		

** $p < .01$.

Experiment 3

In Experiment 3, we examined whether participants who were instructed that the layout was stabilized with respect to their body would stop updating their actual heading with respect to the layout and would, instead, adopt a body-stabilized, egocentric frame of reference.

Method

Participants. Participants were 8 female and 8 male undergraduates at Michigan State University who participated as partial fulfillment of course requirements.

Materials, design, and procedure. The materials, design and procedure were similar to Experiment 1 except for the following two modifications:

1. Prior to the physical turn of the participants during the testing phase, they were given a body-stabilized instruction (e.g., "When you physically turn your body, the objects on the floor will move the same degree as you turn. Hence, after you turn right, you will be still facing the cell phone").

2. We used a new tracking system, InterSense IS-900 (InterSense Incorporated, Bedford, MA), because of an upgrade to the experiment facility. The new tracking system performed identically to the original system except with a considerably increased range and slightly decreased latency, and, thus, is not likely a different factor in these experiments.

Results and Discussion

Pointing accuracy and pointing latency as a function of actual–imagined distance and learning–imagined distance are presented in

Table 5. Means for each participant and each condition were analyzed in repeated-measures ANOVAs in terms of actual–imagined distance (0° and 90°) and learning–imagined heading (0° and 90°). The ANOVA results for pointing accuracy and pointing latency are presented in Table 6. In both angular error and pointing latency, only the main effect of learning–imagined was significant. The results clearly indicate that after being instructed that the objects were arrayed around the body in a body-stabilized display, people used a body-stabilized frame of reference to access the information array.

General Discussion

Current 3D graphics and tracking technology allow designers to display information arrays around a mobile AR user with respect to a body-stabilized or an environment-stabilized frame of reference. There have been no prior studies conducted to investigate which frame of reference mobile users use and what factors may influence the users' frame of reference choices. This study, through the use of the paradigm developed to investigate human spatial memory and spatial updating in physical environments (Mou, McNamara, et al., 2004), suggests that users with no prior experience of mobile AR systems tend to use an environment-stabilized frame of reference to access information arrays presented in AR environments. In other words, people expect the information arrays of virtual objects in AR environments to behave like arrays of objects in physical environments (i.e., when they rotate their body, objects stay in their locations relative to the physical environment). This study also suggests that users who briefly experience the egocentrically centered display of virtual objects or those who are instructed that the display is egocentrically centered are able to quickly adopt a body-stabilized frame of reference to code and access the locations of virtual objects in the physical environment.

Why do naïve users think the locations of the virtual objects are stabilized with respect to the environment? One apparent explanation is that from birth on human beings perceive that the locations of objects in the environment are independent of their own locomotion and, thus, the relationship between their body's locomotion and changes of self-to-object relations are represented in their cognitive system. To efficiently locomote in an environment where objects are not always visible, humans have to develop the ability to update locations of objects in the environment without visual guidance (Farrell & Robertson, 1998; Mou, McNamara, et al., 2004; Presson & Montello, 1994; Rieser, 1989; Rieser et al., 1986; Sholl & Bartels, 2002; Simons & Wang, 1998; Waller et al., 2002; Wang & Simons, 1999). People couple their motions and locomotion with an automatic spatial updating of the representation of object locations. They do so by coupling their locomotion with the perception of change in the spatial relations between the body and objects in the environment during their interaction with the environment (Rieser, 1999; Rieser, Pick, Ashmead, & Garing, 1995). People with no prior experience in mobile AR systems simply interpret the relation between their locomotion and the locations of virtual objects with the mental model they use to interpret the physical world.

On the other hand, the results of Experiments 2 and 3 show that this lifetime of experience with physical objects can be quickly replaced with a model of virtual object arrays that move with the body. In Experiment 2, participants perceived for only 2 min that the locations of virtual objects stayed stationary with respect to their body rotation. Their spatial updating behavior indicated that people in general tend to use body-stabilized frames of reference to code and access the locations of virtual objects after experiencing the behavior of these objects in the new AR layout. This means that people couple their motions and locomotion with a cancelation of the spatial updating of the representation of object locations. They do so during their interaction with the environment by coupling their locomotion with the perception of "unchange" in the spatial relations between the body and objects in the environment.

The quickness with which the participants could adopt the egocentric array of object locations suggests that they may be using prior experience with objects and environments that move with the body. On some occasions, people do perceive that objects move with them in the physical world. For example, objects that are physically attached to the body, such as a wristwatch or a pocket's contents, stay stationary with respect to a user. Information arrays on the traditional mobile information systems such as laptops, PDAs, and cell phones can move with us given that the physical parts of the system stay stationary with respect to the users. Also, when people move via vehicles such as cars or boats, there is a local environment (i.e., the car or boat cabin) that remains fixed relatively to the body but moves relative to the larger external environment such as the road or sea. We speculated that people might have a mental model in favor of a body-stabilized frame of reference that can accommodate arrays of virtual objects that move with the body although they have no visible means of

Table 5
Pointing Latency (in Seconds) and Pointing Accuracy (in Degrees) as a Function of Actual–Imagined (A–I) Distance and Learning–Imagined (L–I) Distance in Experiment 3

	A–I = 0°				A–I = 90°			
	Latency		Accuracy		Latency		Accuracy	
L–I	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
0°	4.225	1.854	19	11	4.283	2.206	20	7
90°	6.725	4.349	30	25	6.481	2.624	37	21

Table 6
Analysis of Variance Results for Pointing Latency and Pointing Accuracy in Actual–Imagined (A–I) and Learning–Imagined (L–I) Conditions in Experiment 3

	$F(1, 15)$		Cohen's f	
Source	Latency	Accuracy	Latency	Accuracy
L–I	18.68**	9.98**	1.12	.82
Error	4.72	327.05		
A–I	.09	2.43	.08	.40
Error	1.51	89.90		
L–I × A–I	.08	1.38	.07	.30
Error	4.62	133.95		

** $p < .01$.

attachment to the body. This consideration was supported by the results of Experiment 3, which showed that even without any direct experience and with only oral instruction that the objects were fixed relative to the body (body-stabilized frame of reference), people were able to use body-stabilized frames of reference to code and access the locations of virtual objects. The results of both Experiments 2 and 3 also suggest that with respect to spatial updating, the cognitive system is highly flexible in interactions with the environment.

Can users of AR systems remember and make use of arrays of 3D objects that move around the body even when they are more than 1 m away from the body? The results of these studies suggest that high quality, mobile AR interfaces may be able to leverage the capacity of human spatial memory and spatial updating mechanisms for efficient access to information items around the body. In this study, we attempted to (a) identify the default frame of reference in coding virtual objects in a high-quality AR mobile system and (b) determine whether experience and oral instruction could alter it. Further studies should investigate how people encode the locations of virtual objects on occasions in which both body-stabilized and environment-stabilized frames of reference are necessary. It remains to be seen whether the updating of these virtual objects interferes with the updating process for objects in the physical environment. This notwithstanding, the current study provides answers to the questions raised in the introduction: Users with no prior experiences in mobile AR systems tend to use environment-stabilized reference frames to encode and access information arrays around their body. Evidently, experiences with or oral instructions of a body-stabilized display allow users to adopt a body-stabilized frame of reference instead.

Moreover, the results of this study imply that the AR system might be an excellent tool in the basic research of human spatial memory and spatial updating because the results (actual–imagined effect and learning–imagined effect) observed in physical environments (Mou, McNamara, et al., 2004) are also observed in the present AR experimental setting. The AR system can also be used to investigate the effect of recoupling perception and action because the relation between the two can be manipulated easily in the AR environments.

References

Azuma, R. (1997). A survey of augmented reality. *Presence, 6,* 355–385.

Barfield, W., & Caudell, T. (2000). *Fundamentals of wearable computers and augmented reality.* Hillsdale, NJ: Erlbaum.

Farrell, M. J., & Robertson, I. H. (1998). Mental rotation and the automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 227–233.

Levine, M., Jankovic, I. H., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General, 111,* 157–175.

Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 162–170.

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 142–157.

Mou, W., Zhang, K., & McNamara, T. P. (2004). Frames of reference in spatial memories acquired from language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 171–180.

Owen, C., Tang, A., & Xiao, F. (2003, October). *ImageTclAR: A blended script and compiled code development system for Augmented Reality*. Paper presented at STARS2003, The International Workshop on Software Technology for Augmented Reality Systems, Tokyo, Japan.

Presson, C. C., & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception, 23,* 1447–1455.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1157–1165.

Rieser, J. J. (1999). Dynamic spatial orientation and the coupling of representation and action. In R. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes* (pp. 168–191). Baltimore: Johns Hopkins University Press.

Rieser, J. J., Guth, D. A., & Hill, E. W. (1986). Sensitivity to perspective structure while walking without vision. *Perception, 15,* 173–188.

Rieser, J. J., Pick, H. L., Ashmead, D. H., & Garing, A. E. (1995). Calibration of human locomotion and models of perceptual-motor organization. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 480–497.

Sholl, M. J., & Bartels, G. P. (2002). The role of self-to-object updating in orientation-free performance on spatial memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 422–436.

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science, 9,* 315–320.

Waller, D., Montello, D. R., Richardson, A. E., & Hegarty, M. (2002). Orientation specificity and spatial updating. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1051–1063.

Wang, R. F., & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition, 70,* 191–210.

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 151–168.

Received May 17, 2004
Revision received July 30, 2004
Accepted August 4, 2004 ■